**Uber Data Analysis**

**Data**

The data contains the details for the Uber rides across various boroughs (subdivisions) of New York City at an hourly level and attributes associated with weather conditions at that time.

* **pickup\_dt**: Date and time of the pick-up.
* **borough**: NYC's borough.
* **pickups**: Number of pickups for the period (hourly).
* **spd**: Wind speed in miles/hour.
* **vsb**: Visibility in miles to the nearest tenth.
* **temp**: Temperature in Fahrenheit.
* **dewp**: Dew point in Fahrenheit.
* **slp**: Sea level pressure.
* **pcp01**: 1-hour liquid precipitation.
* **pcp06**: 6-hour liquid precipitation.
* **pcp24**: 24-hour liquid precipitation.
* **sd**: Snow depth in inches.
* **hday**: Being a holiday (Y) or not (N).

**Basic Steps:**

* + 1. Display the top 5 rows.
    2. Display the last 5 rows
    3. Check the shape of dataset.
    4. Check the datatypes of each feature.
    5. Check the Statistical summary
    6. Check the null values
    7. Check the duplicate values
    8. Check the anomalies or wrong entries.
    9. Check the outliers and their authenticity.
    10. Do the necessary data cleaning steps like dropping duplicates, unnecessary columns, null value imputation, outliers treatment etc.

1. **Pickup Analysis**
   * What is the total number of Uber pickups across all boroughs?
   * Which borough has the highest average number of hourly pickups?
   * How do the number of pickups vary across different hours of the day?
   * Which day of the week has the highest number of pickups?
2. **Weather Impact**
   * What is the correlation between temperature and the number of pickups?
   * How does visibility impact the number of pickups?
   * Is there a relationship between wind speed and the number of pickups?
   * How does precipitation (1-hour, 6-hour, 24-hour) affect the number of pickups?
3. **Seasonal Trends**
   * How do the number of pickups vary across different seasons (winter, spring, summer, fall)?
   * What is the average number of pickups during holidays compared to non-holidays?
   * How does snow depth influence the number of pickups?
4. **Hourly Trends**
   * What are the peak hours for Uber pickups in each borough?
   * How do the number of pickups change during rush hours (e.g., 7-9 AM, 5-7 PM)?
   * What is the average number of pickups during late-night hours (e.g., 12 AM - 4 AM)?
5. **Borough Comparison**
   * How do pickup trends differ between boroughs during different weather conditions?
   * Which borough shows the highest increase in pickups during holidays?
   * How does the number of pickups compare between weekdays and weekends for each borough?
6. **Weather Extremes**
   * How do extreme weather conditions (e.g., very high or very low temperatures, heavy rainfall, snowstorms) affect the number of pickups?
   * What is the impact of visibility less than 1 mile on the number of pickups?
7. **Data Correlations**
   * Is there a correlation between sea level pressure and the number of pickups?
   * How do different weather variables (temperature, dew point, wind speed, visibility) collectively impact the number of pickups?
   * What is the relationship between holiday status and weather conditions on the number of pickups?
8. **Growth Insights**
   * Which weather conditions are most favorable for Uber pickups, and how can this information be used to optimize driver availability?
   * Based on the data, what recommendations can be made to Uber to increase pickups during low-demand periods?

|  |  |  |
| --- | --- | --- |
| **S No.** | **Topic** | **Page No.** |
| 1 | Cover Page | 1 |
| 2 | Index | 2 |
| 3 | List Of Tables | 3 |
| 4 | List of Figures | 3 |
| 5 | Problem Statement | 4 |
| 6 | Data Dictionary | 4 |
| 7 | Basic EDA | 5-15 |
| 8 | Problem | 16-18 |